

---

# 因子分析

因子分析尝试识别出基础变量（或称**因子**）来解释在一组观察到的变量中体现的相关模式。因子分析通常用于数据降维，其目的是识别出少数几个因子来解释大多数在众多显性变量中所观测到的方差。因子分析也可用于生成关于因果机制的假设或过滤变量以用于随后的分析（例如：在执行线性回归分析之前识别共线性）。

因子分析过程提供了高度的灵活性：

- 有 7 种因子抽取的方法。
- 有 5 种旋转方法，包括直接 Oblimin 方法和非正交旋转的最优斜交。
- 有 3 种计算因子得分的方法，并且得分可以另存为变量以进行进一步分析。

**示例。** 什么基础态度使人们回答政治调查上的问题？检查调查项中的相关性显示，项的各种子组有显著的交叉 - 关于税的问题显得彼此相关，关于军事的问题显得彼此相关，等等。使用因子分析，您可以调查基础因子的数量，并且，在许多情况下，还可以识别这些因子在概念上所代表的含义。此外，您可以计算每个响应者的因子得分，然后这些得分可以用于以后的分析。例如，您可以建立 logistic 回归模型以根据因子得分预测投票行为。

**统计。** 对于每个变量：有效个案数、平均值和标准差。对于每项因子分析：变量的相关性矩阵，包括显著性水平、行列式和逆；再生相关性矩阵，包括反映像；初始解（公因子方差、特征值以及解释的方差所占的百分比）；对取样充分性的 Kaiser-Meyer-Olkin 度量以及 Bartlett 球形度检验；未旋转的解，包括因子载荷、公因子方差和特征值；以及旋转解，包括旋转的模式矩阵和转换矩阵。对于斜交旋转：旋转的模式和结构矩阵；因子得分系数矩阵和因子协方差矩阵。图：特征值的碎石图以及前两个或前三个因子的载荷图。

## 因子分析数据注意事项

**数据。** 变量在区间或比率级别应该是定量变量。分类数据（例如：宗教或原产国家/地区）不适合因子分析。可计算 Pearson 相关性系数的数据应该适合于因子分析。

**假设。** 对于每对变量，数据应具有二元正态分布，且观察值应是独立的。因子分析模型指定变量是由公共因子（模型估计的因子）和特殊因子（不在观察到的变量之间交叉）确定的；计算的估计值所基于的假设是所有唯一因子相互之间不相关并与公共因子不相关。

## 获取因子分析

1. 从菜单中选择：

**分析 > 降维 > 因子...**

2. 选择用于因子分析的变量。

---

## 因子分析：选择个案

选择用于分析的个案：

1. 选择一个选择变量。
2. 单击**值**以输入整数作为选择值。

因子分析中仅使用具有该选择变量值的个案。

---

## 因子分析: 描述

**统计。** **一元描述**包括每个变量的平均值、标准差和有效个案数。**初始解**显示初始公因子方差、特征值和已解释方差的百分比。

**相关性矩阵。**可用选项为系数、显著性水平、行列式、KMO 和 Bartlett 球形度检验、逆、再生和反映像。

- **KMO 和 Bartlett 球形度检验。**取样足够度的 Kaiser-Meyer-Olkin 度量检验变量之间的偏相关性是否较小。Bartlett 的球形度检验可检验相关性矩阵是否为恒等矩阵, 该检验可以指示因子模型不适当。
- **再生。**从因子解估计的相关性矩阵。还显示残差(估计相关性和观察相关性之间的差分)。
- **反映像。**反映像相关性矩阵包含偏相关系数的相反数, 而反映像协方差矩阵包含偏协方差的相反数。在一个好的因子模型中, 大部分非对角线的元素将会很小。变量的取样充分性度量显示在反映像相关性矩阵的对角线上。

---

## 因子分析: 抽取

**方法。**使您可以指定因子抽取的方法。可用方法为主成分分析、未加权最小二次方、广义最小二乘法、极大似然、主轴因子分解、Alpha 因子分解和映像因子分解。

- **主成分分析 (Principal Components Analysis).**一种因子抽取方法, 用于形成观察变量的不相关的线性组合。第一个成分具有最大的方差。后面的成分对方差的解释的比例逐渐变小, 它们相互之间均不相关。主成分分析用来获取最初因子解。它可以在相关性矩阵是奇异矩阵时使用。
- **未加权最小二次方法 (Unweighted Least-Squares Method).**一种因子抽取方法, 该方法可以使观察的相关性矩阵和再生的相关性矩阵之间的差的平方值之和最小(忽略对角线)。
- **广义最小二乘法 (Generalized Least-Squares Method).**一种因子抽取方法, 该方法可以使观察的相关性矩阵和再生的相关性矩阵之间的差的平方值之和最小。相关系数要进行加权。权重为他们单值的倒数, 这样单值高的变量, 其权重比单值低的变量的权重小。
- **极大似然法 (Maximum-Likelihood Method).**一种因子抽取方法, 在样本来自多变量正态分布的情况下, 它生成的参数估计最有可能生成了观察到的相关性矩阵。将变量单值的倒数作为权重对相关性的进行加权, 并使用迭代算法。
- **主轴因子分解 (Principal Axis Factoring).**一种从初始相关性矩阵抽取因子的方法, 在初始相关性矩阵中, 多元相关系数的平方放置于对角线上作为公因子方差的初始估计值。这些因子载荷用来估计替换对角线中的旧公因子方差估计值的新的公因子方差。继续迭代, 直到某次迭代和下次迭代之间公因子方差的变化幅度能满足抽取的收敛性条件。
- **$\alpha$ .**一种因子抽取方法, 它将分析中的变量视为来自潜在变量全体的一个样本。此方法使因子的  $\alpha$  可靠性最大。
- **映像因子分解 (Image Factoring).**由 Guttman 开发的因子抽取方法, 它基于映像理论。变量的公共部分(称为偏映像)定义为其对剩余变量的线性回归, 而非假设因子的函数。

**分析。**使您可以指定相关性矩阵或协方差矩阵。

- **相关性矩阵。**在分析中使用不同的刻度测量变量时很有用。
- **协方差矩阵。**当您想将因子分析应用于每个变量具有不同方差的多个组时很有用。

**抽取。**可以保留特征值超过指定值的所有因子, 也可以保留特定数量的因子。

**输出。**使您可以请求未旋转的因子解和特征值的碎石图。

- **未旋转的因子解 (Unrotated Factor Solution).**显示未旋转的因子载荷(因子模式矩阵)、公因子方差和因子解的特征值。

- **碎石图。**与每个因子相关联的方差的图。该图用于确定应保持的因子个数。通常该图显示大因子的陡峭斜率和剩余因子平缓的尾部之间明显的中断（碎石）。

**最大收敛性迭代次数。**使您可以指定算法估计解的过程所采取的最大步骤数。

---

## 因子分析：旋转

**方法。**使您可以选择因子旋转的方法。可用的方法有最大方差、直接 Oblimin、最大四次方值、最大平衡值或最优斜交。

- **最大方差法 (Varimax Method).**一种正交旋转方法，它使得对每个因子有高负载的变量的数目达到最小。该方法简化了因子的解释。
- **直接 Oblimin 方法。**一种斜交（非正交）旋转方法。当 delta 等于 0（缺省值）时，解是最斜交的。delta 负得越厉害，因子的斜交度越低。要覆盖缺省的 delta 值 0，请输入小于等于 0.8 的数。
- **最大四次方值法 (Quartimax Method).**一种旋转方法，它可使得解释每个变量所需的因子最少。该方法简化了观察到的变量的解释。
- **最大平衡值法 (Equamax Method).**一种旋转方法，它是简化因子的最大方差法与简化变量的最大四次方值法的组合。它可以使得高度依赖因子的变量的个数以及解释变量所需的因子的个数最少。
- **最优斜交旋转 (Promax Rotation).**斜交旋转，可使因子相关联。该旋转可比直接最小斜交旋转更快地计算出来，因此适用于大型数据集。

**输出。**使您可以在旋转解上包含输出以及前两个或前三个因子的载荷图。

- **旋转解 (Rotated Solution).**必须选择旋转方法才能获得旋转解。对于正交旋转，会显示已旋转的模式矩阵和因子转换矩阵。对于斜交旋转，会显示模式、结构和因子相关性矩阵。
- **载荷图 (Factor Loading Plot).**前三个因子的三维因子载荷图。对于双因子解，那么显示二维图。如果只抽取了一个因子，那么不显示图。如果要求旋转，那么图会显示旋转解。

**最大收敛性迭代次数。**使您可以指定算法执行旋转所采取的最大步骤数。

---

## 因子分析：得分

**保存为变量。**为最终解中的每个因子创建一个新变量。

**方法。**计算因子得分的可选方法有回归、Bartlett 和 Anderson-Rubin。

- **回归法 (Regression Method).**一种估计因子得分系数的方法。生成的分数的平均值为 0，方差等于估计的因子分数和真正的因子值之间的平方多相关性。即使因子是正交的，分数也可能相关。
- **Bartlett 得分。**一种估计因子得分系数的方法。所产生分数的平均值为 0。使整个变量范围中所有唯一因子的平方和达到最小。
- **Anderson-Rubin 方法 (Anderson-Rubin Method).**一种估计因子得分系数的方；它对 Bartlett 方法做了修正，从而确保被估计的因子的正交性。生成的分数平均值为 0，标准差为 1，且不相关。

**显示因子得分系数矩阵。**显示与变量相乘以获取因子得分的系数。还显示因子得分之间的相关性。

---

## 因子分析：选项

**缺失值。**允许您指定如何处理缺失值。可用选项为按列表排除个案，成对排除个案，或替换为平均值。

**系数显示格式。**使您可以控制输出矩阵的各个方面。按大小对系数进行排序，并排除绝对值小于指定值的系数。

---

## FACTOR 命令的附加功能

使用命令语法语言还可以：

- 在抽取和旋转过程中指定迭代的收敛性标准。
- 指定单独的旋转因子图。
- 指定保存多少因子得分。
- 指定用于主轴因子分解方法的对角线值。
- 将相关性矩阵或因子载荷矩阵写入磁盘以便以后分析。
- 读取和分析相关性矩阵或因子载荷矩阵。

请参阅 *命令语法参考* 以获取完整的语法信息。